Lecture 3 Worksheet

Chrysafis Vogiatzis

Every worksheet will work as follows.

1. You will be asked to form a group with other students in the class: you can make this as big or as small as you'd like, but groups of 4-5 work best.
2. Read through the worksheet, discussing any questions with the other members of your group.
   - You can call me at any time for help!
   - I will also be interrupting you for general guidance and announcements at random points during the class time.
3. Answer each question (preferably in the order provided) to the best of your knowledge.
4. While collaboration between students is highly encouraged and expected, each student has to submit their own version.
5. You will have 24 hours (see gradescope) to submit your work.

Activity 1: Game of dies

Problem 1

Consider a game where you roll two fair dice (i.e., where each number on the side of the dice from 1 to 6 has an equal probability of appearing). What is the probability that the sum of the numbers on the two dice is 7?

Answer to Problem 1.

Problem 2

What is the probability the sum of the numbers on the two dice is 7 given that the first of the two dice rolled on a 3?

Answer to Problem 2.

Problem 3

Based on your answers on part (a) and (b), what can you claim about the independence of the events "the sum of the two dice is equal to 7" and "the first dice is a 3"? Is that true for all pairs of events "the sum of the two dice is 7" and "the first dice is a i" where $i = 1, 2, 3, 4, 5, 6$?

Answer to Problem 3.

Problem 4

Prove or disprove[1] the following statement.

- When throwing two dies, the two events "the sum of the two dies is $j = 2, 3, \ldots 12$" and "the first die is a $i = 1, 2, \ldots, 6$" for $i < j$ are independent events.

[1] To disprove a statement, you may simply find an example where the statement is **not** true.

Answer to Problem 4.

Activity 2: Quality control revisited

Problem 5

A manufacturing facility is making 2 different products. Every product can be classified as defective (D) or non-defective (ND). In addition to that, some products appear to have cosmetic damage (C) or not (NC). The company has collected data for both products over the last 400 items for each. [2]

[2] You may treat these numbers as "probabilities": for example a product 1 is defective and has cosmetic damage with probability 5/400, whereas a product 2 that is known to be defective has cosmetic damage with probability 2/20.

Product 1:

Def.	Cosm. dam. Yes (C)	No (NC)	Total
Yes (D)	5	23	28
No (ND)	24	348	372
Total	29	371	400

Product 2:

Def.	Cosm. dam. C	NC	Total
D	2	18	20
ND	38	342	380
Total	40	360	400

You pick up an item from the recent production of Product 1. If you see it has cosmetic damage, does this alter your perception that the product is defective? [3]

[3] Could we compare the probability that an item is defective (let it be $P(D)$) with the probability that it is defective *given* that it has cosmetic damage (let it be $P(D|C)$)? What if these probabilities are equal to one another? What if they are not?

Answer to Problem 5.

Problem 6

Using the data provided from the previous problem, what can you deduce about Product 2? Does knowing that it has cosmetic damage alter your perception that the product is defective?

Answer to Problem 6.

Activity 3: The birthday problem

Our class has 77 students. If our class had 366 students (assume for now with me that February 29th does *not* exist and, hence, every year has 365 days), then we would be guaranteed that at least two of you share the same birthday.

The question though becomes: in a class of 77, like ours, what is the probability that two of you have the same birthday?

Problem 7

Let's start simple. If there are only two of you, what is the probability that you share the same birthday?

Answer to Problem 7.

Problem 8

With your answer in Problem 7 in mind, add a third person in the mix. What is the probability that at least two people share a birthday in this group of 3? [4]

Answer to Problem 8.

In many instances when interested in finding the probability of an event E, it is easier to calculate $P(\overline{E})$ and use that to find $P(E) = 1 - P(\overline{E})$, rather than trying to calculate $P(E)$ immediately!

[4] Hint: consider the event that you do not have the same birthday as E and then calculate $P(\overline{E})$. Also: how many possible triplets of birthday dates can you create such that no two birthdays are the same? And how many possible triplets of birthday dates can you create in total? Probability can de calculated as the first number over the second on, if only we knew how to count the number of events...

Problem 9

Based on your reasoning in Problems 7 and 8, you must have reached a probability of

$$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \ldots \cdot \frac{365-n+1}{365}$$

for the event that no two people share a birthday in a group of n people. Using the fact that $P(E) = 1 - P(\overline{E})$, we can deduce that the probability that two people share a birthday is:

$$P(\textit{share birthday}) = 1 - \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \ldots \cdot \frac{365-n+1}{365} \qquad (1)$$

What does expression (1) evaluate to in a class of 77 students? What does it evaluate to in a class of 25 students?

Answer to Problem 9.

Next time a person in any class of a significant size shares the same birthday with you, remember that this is decidedly **not** the biggest coincidence in the world, but instead a rather common observance.

Say we had run this for a many values of n, starting from $n = 1$ (one person alone has a 0% chance of sharing the birthday with someone else), $n = 2$ (two people sharing a birthday with a $1/365$ chance), $n = 3$ (triplet of people sharing a birthday), and so on, until $n = 160$ people. We would have then obtained a figure like the one in Figure 1. What do you observe?

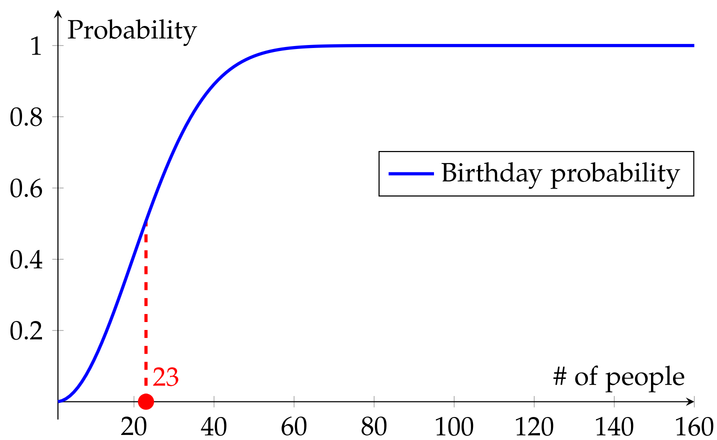

Figure 1: The birthday problem probabilities, visualized. It is at 23 people that this is roughly equal to 50%!

See how few people are needed for the probability to get *almost* equal to 100%! This is a very important realization and helps us understand that sometimes rather common observations are viewed as rare.